Vision Res., Vol. 37, No. 4, pp. 397–406, 1997

Spatial Layout Affects Speed Discrimination

PREETI VERGHESE,*† LELAND S. STONE*

Received 7 March 1996; in revised form 15 May 1996

We address a surprising result in a previous study of speed discrimination with multiple moving gratings: discrimination thresholds decreased when the number of stimuli was increased, but remained unchanged when the area of a single stimulus was increased [Verghese & Stone (1995). *Vision Research, 35*, 2811–2823]. In this study, we manipulated the spatial- and phase relationship between multiple grating patches to determine their effect on speed discrimination thresholds. In a *fusion* experiment, we merged multiple stimulus patches, in stages, into a single patch. Thresholds increased as the patches were brought closer and their phase relationship was adjusted to be consistent with a single patch. Thresholds increased further still as these patches were fused into a single patch. In a *fission* experiment, we divided a single large patch into multiple patches by superimposing a cross with luminance equal to that of the background. Thresholds decreased as the large patch was divided into quadrants and decreased further as the quadrants were maximally separated. However, when the cross luminance was darker than the background, it was perceived as an occluder and thresholds, on average, were unchanged from that for the single large patch. A control experiment shows that the observed trend in discrimination thresholds is not due to the differences in perceived speed of the stimuli. These results suggest that the parsing of the visual image into entities affects the combination of speed information across space, and that each discrete entity effectively provides a single independent estimate of speed. Copyright © 1997 Published by Elsevier Science Ltd

Segmentation Speed perception Temporal frequency Perceived speed Occlusion Motion Area MT

INTRODUCTION

In the last 40 years, a large part of the vision community has concentrated on early visual processing and, in particular, has focused on the decomposition of the image into functional components. At present, there is considerable psychophysical and physiological evidence that the visual system first decomposes the image using arrays of local mechanisms tuned for specific stimulus properties such as spatial frequency, orientation, direction of motion, etc. (De Valois & De Valois, 1990). However, there is still not a clear understanding of how all of these different local components interact or combine to synthesize a perceptually seamless visual scene. The dominant view has been that the various modules analyze elementary aspects of the visual stimulus in parallel, and that the interpretation of the visual image into surfaces and objects takes place, subsequently, at a higher level.

In particular, mechanisms selective for motion direction appear as early as striate cortex (V1) in primates (Hubel & Wiesel, 1968), with the middle temporal area (MT) being specialized for motion processing (Zeki, 1974; Maunsell & Van Essen, 1983). Albright (1984) has shown that area MT has a systematic columnar arrangement of neurons selective for different directions of motion, analogous to the orientation columns of striate cortex (Hubel & Wiesel, 1968). In addition, Newsome and co-workers have used lesion studies, electrical stimulation, and single-unit recording to show that MT is critically involved in direction perception (Newsome & Pare, 1988; Newsome *et al.*, 1989; Salzman *et al.*, 1990). Psychophysical evidence also supports direction processing as occurring early, as contrast thresholds for the detection of motion are generally equal to those for discriminating the direction of motion (Watson *et al.*, 1980; Pasternak & Merigan, 1994).

Although the above studies support the existence of direction-selective mechanisms that occur early, He and Nakayama (1994), have shown that direction perception is influenced by "higher-level" representations such as surfaces. They showed that even when the binocular disparity of moving stimuli was left unchanged, the perceived direction of motion depended on whether the stimuli appeared to lie on a single surface or on multiple surfaces. A study by Stoner and Albright (1992) is also consistent with image segmentation affecting motion perception. Their data showed that the response of pattern cells (c.f. Movshon *et al.*, 1985; Rodman & Albright,

*Flight Management and Human Factors Division, NASA Ames Research Center, M/S 262-2 Moffett Field, CA 94035-1000, U.S.A.

†To whom all correspondence should be addressed at present address: Smith Kettlewell Eye Research Institute, 2232 Webster St, San Francisco, CA 94115, U.S.A. [*Fax* +1-415-561-1610; *Email* preeti@skivs.ski.org].

1989) in visual area MT depends on whether a plaid pattern created by the superposition of two orthogonally oriented square-wave gratings, appears to cohere or not. They manipulated coherence by varying the luminance of the plaid intersections. The fact that the luminance of the intersections determined both perceptual coherence and neuronal responses suggests that image parsing affects the integration of direction signals as early as MT.

Speed perception is also thought to be mediated by local mechanisms tuned for speed. Maunsell and Van Essen (1983) demonstrated that a majority of the neurons in area MT have, in addition to direction selectivity, a selectivity for stimulus speed. Movshon and co-workers (Movshon *et al.*, 1988) have shown that the spatial and temporal frequency tuning of at least some MT cells covary, providing true speed selectivity across a range of temporal and spatial frequencies. In addition, Pasternak and Merigan (1994) have shown that bilateral lesions of MT (which included the adjoining medial superior temporal area, MST) caused permanent deficits in speed discrimination thresholds in monkeys. Psychophysical measurements have shown that humans can discriminate speed differences as small as 5%, which are equal to or lower than thresholds for temporal frequency discrimination (McKee *et al.*, 1986). These studies together suggest that speed perception is subserved by specialized speed mechanisms, perhaps as early as MT.

In this study, we explore the effect of image parsing on speed perception. The motivation comes from recent work (Verghese & Stone, 1995) that measured human ability to combine speed information from multiple moving grating patches. Observers were presented with a 2-interval forced-choice (2IFC) paradigm, with *n* grating patches in each interval all moving at the same speed, and were asked to choose the interval with the faster gratings. While increasing the number of stimuli improved speed discrimination, a surprising finding from this study was that increasing the area of a single patch caused no such improvement. Figure 1 replots these data as thresholds (normalized to that for a single, regular-sized grating and averaged across four observers) as a function of total stimulus area. The solid symbols show that the speed difference required to reach threshold decreased as the number of gratings increased. However, this improvement was not due to increased stimulus area *per se*, as thresholds were unchanged when the area of a single grating was increased by a factor of 2, 4 or 6 (open symbols in Fig. 1).

This lack of spatial summation is surprising, given that summation has been reported for detection at threshold contrasts* (Robson & Graham, 1981) and for the detection of correlated dot motion among uncorrelated noise at suprathreshold contrasts (Downing & Movshon, 1989). In general, the increased detectability of a

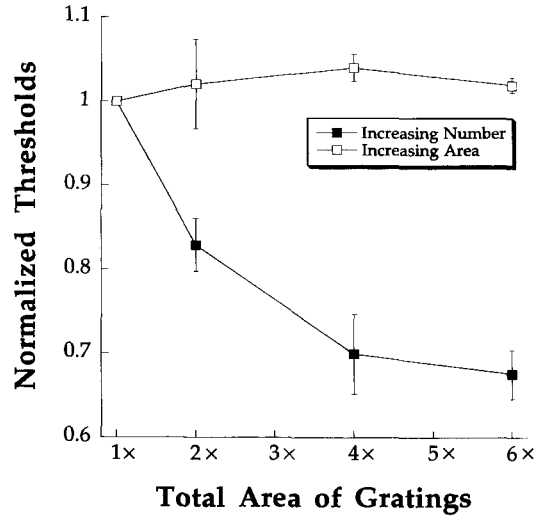

FIGURE 1. Speed discrimination thresholds vs total stimulus area. Thresholds were normalized to that for a single, regular-sized grating and averaged across four observers. Thresholds decreased with increasing grating number as shown by the solid symbols. Thresholds were unchanged with increased grating area as shown by the open symbols. Error bars show standard error across four observers.

stimulus with increasing area, is well modeled by linear summation for small stimuli within a single detector, and probability summation across space between these local, independent detectors for extended stimuli. Therefore, it is surprising that speed perception, which is thought to be mediated by local, speed-selective units, does not show classic area summation.

To explain this intriguing finding, we propose that the decrease in speed discrimination thresholds in the case of multiple gratings is due to each grating patch providing an independent sample of speed, while one large grating gives only a single estimate. If the speed estimates from each patch are independent samples from a noisy distribution, then in the case of multiple patches, averaging across patches improves the estimate of speed. However, in the case of a single large grating, there is only a single noisy sample regardless of the size of the grating, and consequently, no improvement in the precision of the speed estimate with size.

To test this hypothesis, we used two approaches. We measured speed discrimination in a *fusion* experiment, where we merged multiple gratings until they became a single stimulus. We also measured thresholds in a *fission* experiment, where a single large grating was divided into multiple stimuli. Our hypothesis predicts that thresholds will increase in going from multiple gratings to a single one, regardless of how this is achieved. A preliminary report of a subset of these results has appeared elsewhere (Verghese & Stone, 1996).

METHODS

The general method was similar to Verghese and Stone (1995). Our experimental paradigm was a 2IFC speed discrimination task. The stimuli were patches of drifting vertical grating, windowed by a stationary Gaussian

*Our main experiments were conducted at suprathreshold contrasts. In control experiments with contrasts close to detection threshold, increasing the area of a single grating did give the classic summation result.

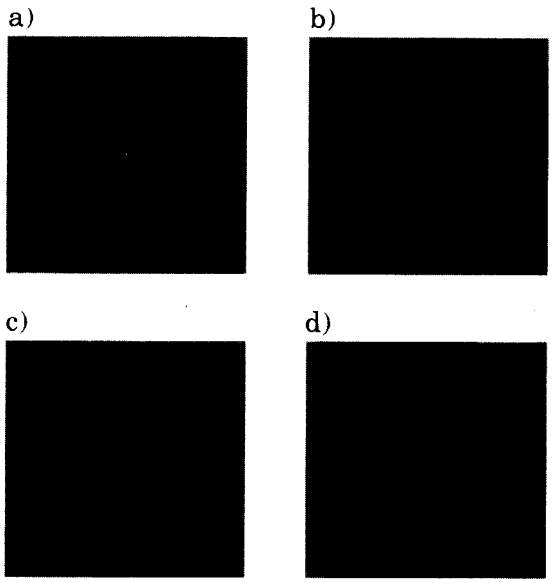

FIGURE 2. Stimuli in the fusion experiment: (a) 3 regular: three grating patches, maximally separated, i.e., with an angular separation of 120 deg; (b) 3 in-phase: three grating patches, with an angular separation 30 deg, and a phase relationship consistent with Gaussian windows on a full-field grating; (c) banana: three stimuli fused into a single banana-shaped patch; (d) 1 large: a single large, circular grating, 3× the area of a regular grating.

centered at an eccentricity of 4 deg from fixation. All the grating patches in each interval moved in the same direction, which was randomly picked to be leftward or rightward. The spatial frequency of the gratings was fixed at 1.5 c/deg, while the temporal frequency of the reference was jittered (± 1 Hz) about 8 Hz. The background luminance was 20 cd/m^2. Unless otherwise stated, the contrast of the gratings was 20%. A trial consisted of two 195 msec intervals, with an interstimulus interval of 500 msec. Each interval typically had the same number of patches. One of the intervals, picked randomly, had all of its grating patches moving at the reference speed, and the other had all its patches moving at a faster test speed. The test speed was picked by a 3-up–1-down staircase. Observers were asked to select the interval with the faster patches. Each condition was run separately in a block of trials, and the conditions that comprised an experiment were run in a single experimental session. At least four sessions were run, resulting in four repetitions of each condition. Thresholds for each condition were obtained from a psychometric function of proportion correct vs speed difference. The raw data for each condition were pooled over all runs and fit with a Weibull function, using a weighted χ^2 minimization. The weights were computed from the number of trials and the proportion correct, assuming a binomial distribution. The threshold was determined from this fit and taken as the speed difference corresponding to 82% correct. All observers in the fission and fusion experiments, except LS, were provided with feedback at the end of each trial.

In the fusion experiment, we used the four stimulus conditions depicted in Fig. 2(a–d), respectively. Figure 2(a) shows the *3 regular* condition, in which three regular-sized patches were maximally separated, i.e., the angular separation of their centers was 120 deg. Each of these patches was windowed by a two-dimensional Gaussian with a spatial spread, $\sigma = 0.4$ deg. Figure 2(b) shows the *in-phase* condition, in which the three patches were brought closer, with a center-to-center separation of 30 deg. Furthermore, their phase relationship was consistent with windows on a full-field grating. In the *banana* condition, the three patches were fused to form a single banana-shaped patch [Fig. 2(c)]. This stimulus was one-quarter of an annulus with a mean radius of 4 deg from fixation. The annulus width was equal to the diameter of a regular patch, and its ends were capped with appropriate halves of the regular patch, in matching phase. The edges of this stimulus were windowed with exactly the same spatial profile as the *regular* Gaussian. The last condition in this experiment was a single *large Gabor patch*, three times the area of a regular patch, with a spatial spread $\sigma = 0.7$ deg [Fig. 2(d)]. The total bounding contour length of the stimuli in the first three conditions of this experiment were kept roughly constant. The stimuli were presented in one of four random configurations on each trial: the one shown, and others rotated by 90, 180 or 270 deg.*

In the fission experiments, we used four stimulus conditions that are depicted in Fig. 3(a–d), respectively. In the *large patch* condition, we measured thresholds for a single large grating windowed by a Gaussian with a spread of 1 deg [Fig. 3(a)]. In the *occluder* condition, we superimposed a cross over the large Gabor patch. The luminance of the cross was darker than the mean luminance (contrast 10%) [Fig. 3(b)]. The cross had a luminance profile that was constant for 0.9 deg, with Gaussian edges ($\sigma = 0.2$ deg) that tapered smoothly to mean luminance. In the *divided* condition, we divided the patch into four quadrants by superimposing the same cross, but with a luminance equal to the background, thus introducing horizontal and vertical gaps across the patch [Fig. 3(c)].† In the *4 regular* condition, four gratings were maximally separated [Fig. 3(d)] and their area and peak contrast were adjusted to be equal to that of the occluded and divided stimuli.

In the occluder condition, to avoid the abrupt onset of a large dark cross in the periphery, we placed dark crosses in all four potential stimulus locations for the entire duration of this condition. The divided condition had an equivalent spatial configuration, although we did not

*For the banana condition, we were concerned that thresholds for the horizontally oriented stimulus might be different than that for the vertically oriented stimulus, as there were many more cycles of the grating orthogonal to the orientation of the window in the former case. However, a comparison of speed discrimination thresholds for horizontally and vertically oriented banana stimuli revealed no difference. We therefore treated the banana stimulus like the other stimuli, and placed it in one of four locations.

†In preliminary experiments, we tested gap sizes of 0.13, 0.5, 0.8 and 1.1 deg. Two of three observers showed similar thresholds for the three larger gaps, while the third observer had thresholds that decreased systematically with increasing gap size.

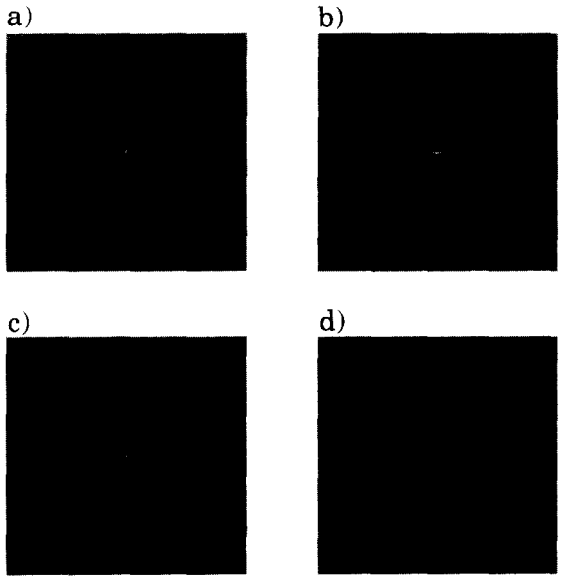

FIGURE 3. Stimuli in the fission experiment. (a) 1 large: a single, large patch; (b) occluded: a cross superimposed on the patch, with contrast 10% darker than the background. The cross had Gaussian edges with a spatial spread of 0.2 deg; (c) divided: same as occluded except that the cross had the same luminance as the background, thus creating the appearance of horizontal and vertical gaps superimposed on the stimulus; (d) 4 regular: four patches, maximally separated, with the same area and peak contrast as the visible grating patches in (b) and (c).

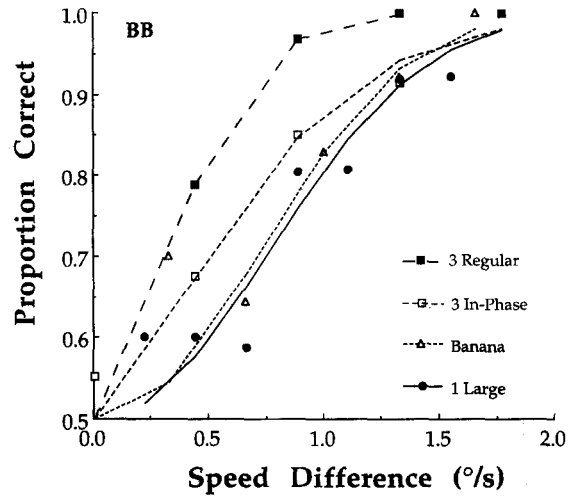

FIGURE 4. Raw data for observer BB in the fusion experiment. Percent correct is plotted vs speed difference and the different symbols represent the data for the different conditions. The smooth lines are Weibull fits to the data. Thresholds were estimated as the speed difference required for 82% correct. The average reduced χ^2 of the fits was 0.46.

physically have to place crosses, as they were equiluminant with the background.

To control for the possibility that differences in perceived speed might have caused the observed differences in speed discrimination, we performed a speed-matching experiment with the stimuli in the fission experiment. Once again, observers were asked to choose the interval with the faster stimulus in a 2IFC task. However, the stimuli in the two intervals were different. The reference, the single large Gabor patch, was assigned randomly to one interval. The test, either the divided Gabor, the occluded Gabor, or four separated Gabor patches, was assigned to the other interval. For each reference test pair, we fixed the speed of the reference and varied the speed of the test according to a simple 1-up–1-down staircase. No feedback was provided for this experiment. We fit the data with a cumulative Gaussian and defined the bias (point of perceived equal speed) as the mean of the underlying Gaussian.

A total of six observers with normal or corrected-to-normal acuity participated in both the fusion and fission experiments, three of whom were naïve as to the purpose of the experiment (ET, CN, and TV). A subset of these observers participated in the control experiments.

RESULTS

Fusion

Figure 4 plots raw data for the fusion experiment for observer BB. Proportion correct is plotted vs speed difference and the smooth lines are Weibull fits to the data for each condition. The psychometric curves move to the right as multiple stimuli are fused into one, with speed discrimination thresholds increasing from the 3 regular condition, to the 3 in-phase, to the banana, to the 1 large Gabor. Figure 5(a) plots speed discrimination thresholds for the different spacing conditions for the six observers. The open symbols show data for our naïve observers, while the solid symbols represent the data of observers wise to the aim of the experiments. Despite the variation in absolute thresholds among observers, five of six* observers have thresholds that increase as the multiple gratings are fused into a single large patch. Thresholds increase as the stimuli are brought closer and in consistent phase, and increase further when they are fused into the banana configuration. They increase further still going from the banana configuration to the large patch. This trend is clearer in Fig. 5(b), which replots the data in Fig. 5(a) with each observer's threshold normalized to that for the 3 regular condition.

In going from the 3 regular condition to the 3 in-phase condition, we changed both the proximity and the phase relationship of adjoining patches. To determine whether proximity alone had an effect, we repeated the first two conditions of the fusion experiment along with an extra condition in which we placed the three gratings in close proximity (an angular separation of 30 deg as in the 3 in-phase condition), but allowed the patches to have a random phase relationship. Figure 6 shows data for this experiment. The thresholds for the new 30 deg, random-phase condition were intermediate between the maximally separated and in-phase conditions for observers CN and PV. Observer ET's threshold in the 30 deg random phase condition appears similar to that in the maximally separated condition.

*Observer LS' data are different from those of others. One possibility is that he is highly practised, probably overly so. The other is that he received no feedback.

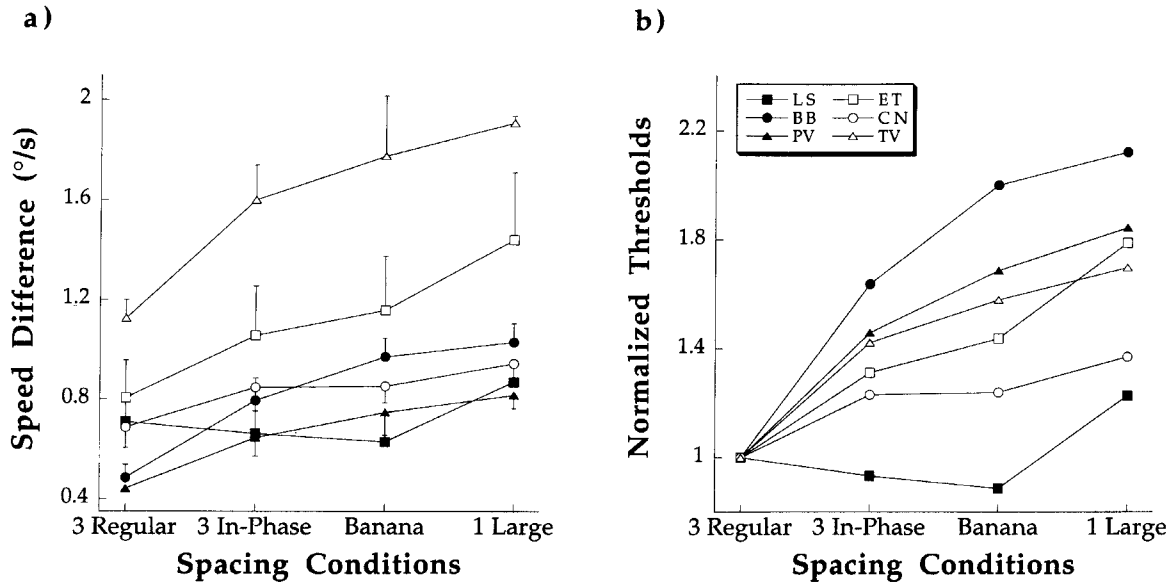

FIGURE 5. Speed discrimination thresholds in the fusion experiment. The different symbols represent data for six different observers. (a) The absolute thresholds are plotted vs spacing conditions. (b) Thresholds normalized to that for the 3 regular condition are plotted vs spacing conditions. In this and subsequent figures, the error bars on individual observer data points represent ± 1 standard deviation of the estimated threshold.

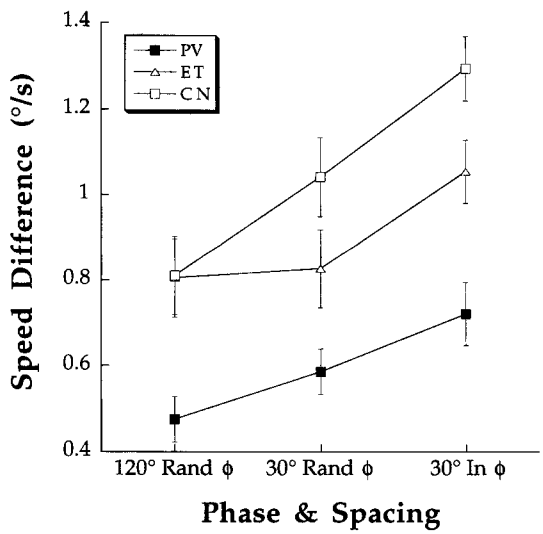

FIGURE 6. Speed discrimination thresholds as a function of phase and spacing for three observers. The first and third conditions correspond to the 3 regular and 3 in-phase conditions of the fusion experiment. In the second condition, the patches had random phase and an angular separation of 30 deg.

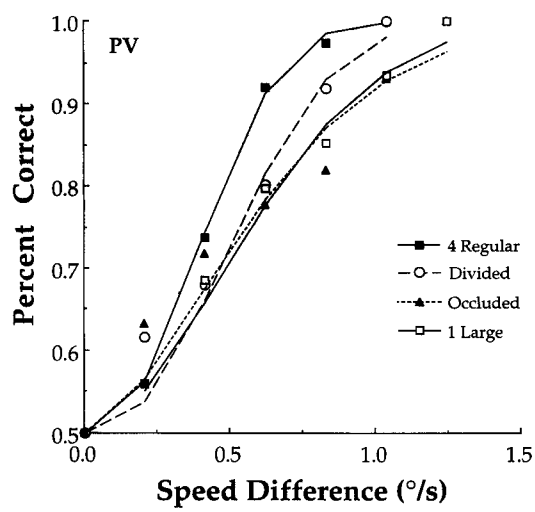

FIGURE 7. Raw data for observer PV in the fission experiment. The different symbols are for the different spacing conditions. The smooth lines are Weibull fits to the data. The average reduced χ^2 of the fits was 0.49.

Fission

Figure 7 plots raw data for the fission experiment for observer PV. Proportion correct is plotted as a function of speed difference and the smooth lines are Weibull fits to the data. Starting with the large Gabor, the psychometric curves move to the left as the large patch was divided and the parts separated, with speed discrimination thresholds decreasing from the large Gabor to the divided Gabor, and decreasing further for four separate Gabors (4 regular). Figure 8(a) plots speed discrimination thresholds for the different conditions in the fission experiment for six observers. This figure uses the same format for spatial configuration on the abscissa as Fig. 5, with the data for the multiple separated patches on the left and the large patch on the extreme right. Four out of six observers had thresholds that were essentially unchanged in going from the large Gabor to the occluded Gabor. All six observers showed a decrease in thresholds in going from the occluded to the divided condition, although the spatial layout of the moving portions of these stimuli were identical. Thresholds for all six observers decreased further when the four patches were maximally separated. This trend is clearer in Fig. 8(b), which replots the data in Fig. 8(a), with each observer's threshold normalized to that for the 4 regular condition.

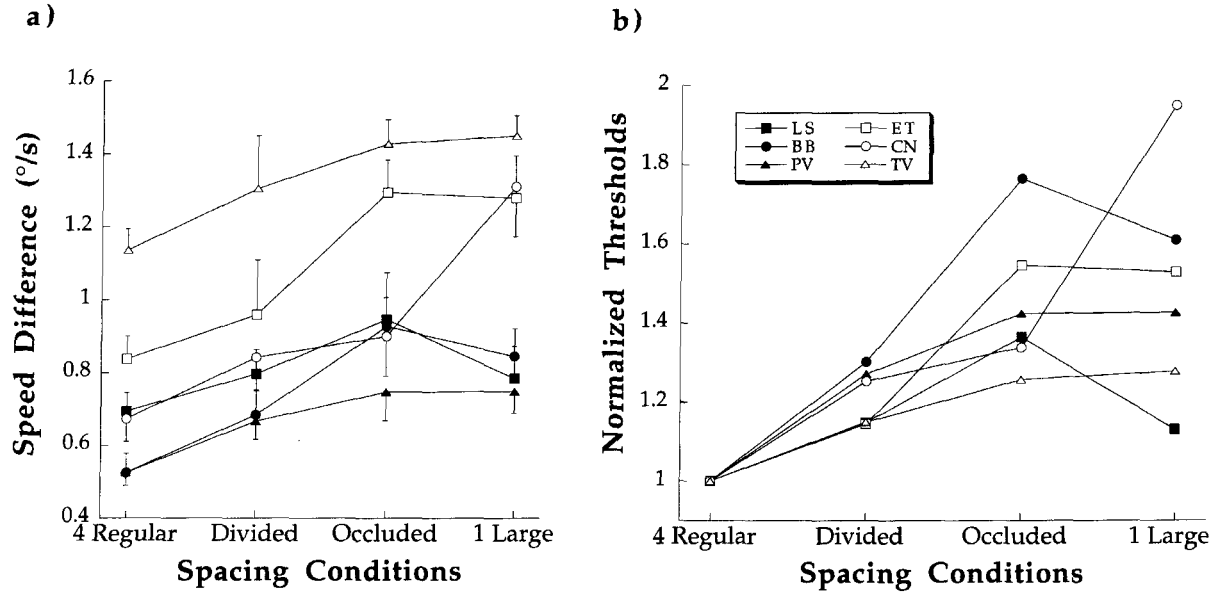

FIGURE 8. Speed discrimination thresholds in the fission experiment. The different symbols represent data for six different observers. (a) The absolute thresholds are plotted vs spacing conditions. (b) Thresholds normalized to that for the 4 regular condition vs spacing conditions.

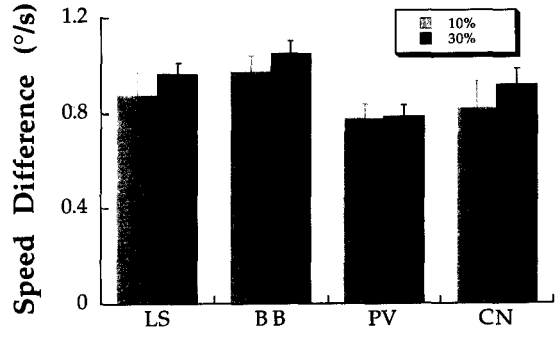

FIGURE 9. Speed discrimination thresholds for the occluded condition in the fission experiment with cross contrasts of 10 and 30%, respectively.

We also performed a control experiment to determine whether the increase in threshold from the divided to the occluded patch was due simply to the increased contrast of the cross in the occluded condition. We compared two conditions in this experiment: one in which the cross had a contrast of 10% (as in the fission experiment), and the other in which the cross had a contrast of 30%. Figure 9 plots the thresholds of four observers for these two values of cross contrast. This increase in contrast caused an increase in average threshold of only 8%, which was not significant across the four observers tested (paired one-tailed t-test; $P = 0.41$). However, the same four observers (including one naïve) showed a significant increase (19% on average) from the divided to the occluded condition ($P < 0.008$).

In related experiments, we used various other means to divide the large patch into quadrants. Superimposing a pair of dark vertical and horizontal lines, as in a cross-hair, on the large patch caused no change in threshold for the two observers tested. More interestingly, thresholds were unchanged when each quadrant of the patch had a different spatial frequency (Fig. 10). To ensure that the quadrants would not be processed by the same spatial frequency channel, we used four spatial frequencies with frequency ratios of 0.5:0.67:1:2, relative to a reference of 1.5 c/deg. Each spatial frequency was randomly allocated to a quadrant and the temporal frequency of each quadrant was adjusted so that all the quadrants had the same speed. Figure 10 compares thresholds for the spatial frequency grid to the single large Gabor. Thresholds were indistinguishable in the two conditions, at least for three of four observers. Observers reported that this multi-spatial frequency patch had a subjective appearance of a single patch of various textures moving at a uniform speed. This, and the fact that most observers had speed discrimination thresholds that were indistinguishable in the spatial frequency grid and large patch conditions, is consistent with both stimuli being perceived as single entities.

Perceived speed

The reference temporal frequency used in our experiment (8 Hz) is such that steep increases in Weber fraction have been observed for small increments above this value (McKee *et al.*, 1986). If the large stimuli used in the fission experiment were perceived to move faster than the four regular patches, it might explain why their thresholds were higher. To test if the basis of the variation in speed discrimination thresholds with spatial layout might be due to increases in perceived speed (or temporal frequency), we measured speed matching between the large Gabor patch (reference) and the other stimuli used in the fission experiment.

Figure 11 plots the ratio of the reference to the test speed vs spacing conditions, for four observers. The test speed was the speed of the test required to match that of the reference. On average, each of the test stimuli

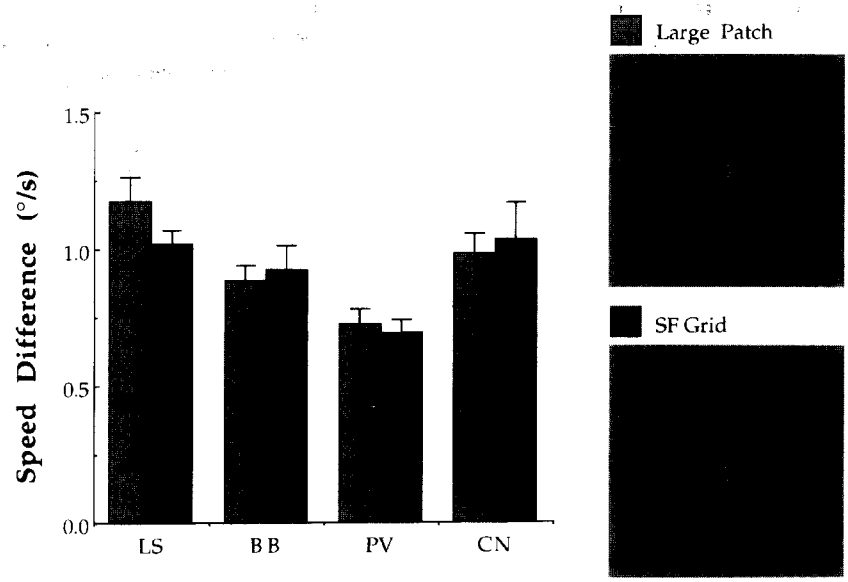

FIGURE 10. Speed discrimination thresholds for the large patch and spatial frequency grid conditions. The large patch is the same as the one used in the fission experiment. The spatial frequency grid was composed of four quadrants, with each quadrant randomly assigned one of four spatial frequencies, with a frequency ratio of 0.5:0.67:1:2. The temporal frequency was adjusted so that all the quadrants moved at the same speed. The velocity of the reference grating was 5.3 deg/sec.

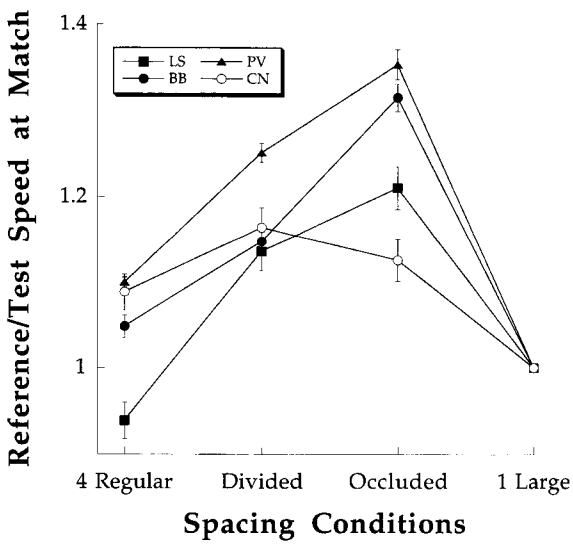

FIGURE 11. Ratio of reference speed to test speed vs spacing conditions in the speed matching experiment. Abscissa values larger than 1 indicate that the test was perceived as faster than the reference. The different symbols represent data for four different observers. The conditions correspond to matches of the 1 large patch used as the reference, to the other stimuli in the fission experiment (see Fig. 3).

appeared faster than the large standard patch, i.e., the speeds of the reference and test were perceived to be equal when the reference actually moved faster than the test. The occluded patch was perceived to move the fastest relative to the reference, followed by the divided patch and finally the four separate patches. As the perceived speed of the four separate patches is comparable to that of the single large patch, a difference in perceived speed cannot account for the large decrease in speed discrimination thresholds in going from one to four patches. Furthermore, the trend in perceived speed is quite different from that of discrimination thresholds in the fission experiment (compare Figs 8 and 11). It is, therefore, unlikely that the observed pattern of thresholds is due to differences in perceived speed.

DISCUSSION

Data from both the fusion and fission experiments are consistent with multiple patches being better than one for speed discrimination. Speed discrimination declines as multiple patches are brought closer and in consistent phase, and declines further when the multiple patches are fused into a single patch. Conversely, speed discrimination improves when a single large stimulus is divided into parts and improves further when each of these parts is maximally separated. There are several experimental factors that may contribute to these effects, such as eye movements, spatial uncertainty, variation of perceived speed with eccentricity, total bounding contour length and motion boundaries, proximity and relative phase, and occlusion. We consider each of these in turn.

Potential factors

Eye movements. In all our experiments, display duration was kept brief (195 msec) to minimize eye movements. If eye movements occurred more often when stimuli were presented at a single location rather than at multiple locations simultaneously, it might account for the increased thresholds in the single-stimulus case. Control experiments that varied the spacing of two regular Gabor patches (Verghese & Stone, 1995) addressed this particular issue. These experiments showed no difference in threshold if the two patches were placed near each other in one location, or in diametrically opposite positions, indicating that the

contamination from eye movements, if it occurs, is probably similar in the single and multiple location cases.

Position uncertainty. It can be argued that the decrease in threshold with increasing number of patches was due to the observer being uncertain about grating position, and that increasing the number of patches increased the probability that a patch appeared in a location that the observer was monitoring. To address this issue, we examined the role of position uncertainty (Verghese & Stone, 1995). Instead of randomizing the position of the gratings, the gratings were presented in fixed positions, thus eliminating position uncertainty. Under these circumstances thresholds as a function of the number of patches decreased at the same rate as in the conditions with random positions, indicating that the decrease in thresholds from one to multiple gratings is independent of position uncertainty.

Bounding contour length and motion boundaries. The contribution of these factors is difficult to isolate, as they sometimes covaried with other factors in our experiments. In the original experiment (Fig. 1), increasing bounding contour length while keeping total stimulus area constant, by going from a single large patch to n patches, each of $1/n$th the area, caused thresholds to decrease. However, increasing the area (and circumference) of a single patch caused no change in thresholds, suggesting that total bounding contour length is not a critical factor. In the fusion experiment, thresholds increased from the 3 regular to the 3 in-phase and further still with the banana configuration (Fig. 5), even though bounding contour length for these stimuli was fixed. This shows that a factor other than bounding contour is responsible for the changes in threshold observed in the fusion experiments (see *Proximity* below). In the fission experiment, increasing contour length by superimposing a cross has two different effects, depending on the luminance of the cross. When the cross is equiluminant with the background, as in the divided case, thresholds decrease. However, when the cross is darker than the background, thresholds remain unchanged for a majority of observers. One possible explanation for the similar thresholds in the occluded and large patch case is that perceived contour length in the occluded case is unchanged, as the contours introduced by the dark cross are assigned to the occluding cross and not to the grating patch (Nakayama *et al.*, 1996). Although our data do not show a consistent effect of contour length or motion boundaries, we cannot rule out some contribution to speed discrimination.

Eccentricity. The lack of improvement in thresholds with increasing area could have been due to the variation in perceived speed with eccentricity. We estimated the variation in perceived speed due to the spatial spread of the large Gabor using data from Johnston and Wright (1986). We assumed that the large patch was sampled by 6 units, each the size of a regular-sized patch, and that each of these units provided an independent speed estimate from an underlying normal distribution of speed. The means of these distributions varied with eccentricity as per Johnston and Wright, and the standard deviations were proportional to the mean (Graham *et al.*, 1987). These samples were averaged (the optimal combination rule in the case when the speed estimates are independent samples from a single normal distribution of speed, see Verghese & Stone, 1995) to obtain a speed estimate in each interval. The model psychometric function obtained by picking the interval with the larger speed accounts for less than 6% of the average threshold increase observed in going from six regular-sized Gabor patches to a single large patch.

Proximity and relative phase. Increasing the proximity of the multiple grating patches and making their phase relationship consistent with a single object causes speed discrimination thresholds to increase in both sets of experiments. The effect of proximity is consistent with a large part of our data: thresholds increased as multiple gratings were merged, decreased when a large grating was divided, and decreased further when the parts were separated. In addition, changing the relative phase of three closely spaced gratings from random to in-phase (i.e., consistent with a single large grating), also caused thresholds to increase.

We showed previously that proximity had no significant effect on thresholds when the angular separation of *two* grating patches, with random phase, was varied between 30 and 180 deg (Fig. 6 of Verghese & Stone, 1995). However, in the present study, decreasing the spacing of *three* patches, from 120 to 30 deg, caused thresholds to increase in two of three observers (Fig. 6). This discrepancy suggests that proximity alone may not be the critical factor. Perhaps, the juxtaposition of three patches is treated more as a single entity than that of two patches.

Occlusion. The effect of the superimposed cross on speed discrimination thresholds clearly depends on the luminance of the cross. When the same cross is equiluminant with the background, the patch appears divided into four quadrants and thresholds are significantly lower than for the single large patch. When the cross luminance is consistent with an occluder and the stimulus is perceived as a single patch moving behind an occluder, thresholds are largely unchanged compared to the single large grating. This threshold increase from the divided to the occluded condition occurs despite the fact that these conditions have identical spatial configurations and identical moving regions. In addition, thresholds do not increase significantly when the contrast of the superimposed cross is increased 3-fold, thus demonstrating that the increase in threshold from the divided to the occluded case is unlikely to be due to contrast masking.

Of all the above factors that we have considered, it appears that only proximity, phase, and occlusion have a clear effect on speed discrimination thresholds. This, together with our previous finding (Verghese & Stone, 1995) that the threshold improvement with multiple patches is due to each patch effectively providing an independent sample of speed, allows a more general conclusion—that speed perception is affected by the

parsing of the visual image into entities, and that each entity effectively provides a single speed estimate. It appears that the multiple speed estimates become inaccessible when the multiple patches are fused into a single patch. This is similar to the result that observers are unable to access component speed when viewing coherently moving plaids (Welch, 1989; Welch & Bowne, 1990).

Relationship to earlier studies

The fact that the thresholds in our experiment remain unchanged in the presence of partial occlusion (Fig. 8) is consistent with He & Nakayama (1992; 1994) who advocate an early parsing of the image into surfaces—prior to the perception of textures, features, or even apparent motion. Our data extend this view by showing that the parsing of multiple patches, even on a single surface, affects speed perception. Furthermore, at a physiological level, there is mounting evidence that neurons in early visual processing areas such as V1 (Zipser, 1994) and MT (Albright & Stoner, 1995) are sensitive to image segmentation cues located outside their classical receptive field. Our data provide complementary psychophysical evidence for image segmentation affecting early visual processing.

Our results with the spatial frequency grid are consistent with the data of Bravo and Watamaniuk (1995). They used a display with dots moving at two very different speeds and asked observers to discriminate changes in the slower speed. Their observers reported seeing two surfaces, one moving at the fast speed and the other at the slow speed, and their ability to discriminate changes around the slower speed was similar to the case when the slow speed was presented alone. Furthermore, speed discrimination thresholds (and the percept of two segmented surfaces) remained unchanged even when the two speeds were generated by a combination of spatial and temporal displacements. This experiment is similar to ours with multiple spatial frequencies: speed discrimination thresholds were unchanged when the large patch was made up of four quadrants, each with a different spatial frequency and a corresponding temporal frequency adjusted so that the grating patches in all the quadrants moved at the same speed (see Fig. 9). The fact that large speed differences result in the percept of segregated surfaces, as well as the fact that speed discrimination is unchanged despite changes in temporal and spatial properties of the stimulus, suggest that speed information itself might determine whether the stimulus is grouped into a single entity or segregated into multiple moving entities.

Underlying mechanisms

What are the mechanisms that might cause speed discrimination to deteriorate as multiple patches are merged? We suggest the two mechanisms that may act early in visual processing: correlation and inhibition.

Correlation. When multiple gratings patches are maximally separated, the improvement in threshold with number is consistent with each patch providing an independent estimate of speed (Verghese & Stone, 1995). In this case, spatially combining samples with independent noise would produce an average signal with a variance smaller than that of each sample. On the other hand, if the correlation of the noise in the speed estimates increases as the patches are brought closer together (e.g., due to a local, common noise source) then, combining these non-independent samples would not decrease the variance of the average signal. This would explain the lack of improvement of threshold with increase in area of a single patch. This is similar to the explanation proposed by Britten *et al.* (1992) for how the outputs of multiple primate MT cells might be pooled to generate perceptual motion thresholds. The fact that psychophysical thresholds were only a factor of 2 or 3 lower than simultaneously measured neuronal thresholds, led these authors to suggest that the neuronal pool that contributes to the perceptual decision is either made up of few independent neurons, or more likely, a large number of neurons whose outputs are significantly correlated. Our results are also consistent with recent studies (Gray *et al.*, 1989; Engel *et al.*, 1991) that found that the firing rates of cortical neurons are more correlated when the neurons are stimulated by a single extended stimulus than when the stimulus is divided into two separate parts.

Correlation may also explain why thresholds in the fusion experiment appear slightly higher in the case of the single large grating, than the banana. Although both are single stimuli, the banana stimulus is more extended. If correlation increases with proximity, then the multiple samples from the banana patch would be less correlated than the more spatially compact circular patch.

Inhibition. We assume a standard visual front-end with multiple arrays of receptors at each location, each tuned to a specific stimulus attribute. If there are inhibitory connections between neighboring units tuned to the same speed (or temporal frequency), then increasing the area of a stimulus could reduce the response of units stimulated by the extended stimulus. Inhibition could be either subtractive (e.g. Allman *et al.*, 1985) or divisive (e.g. Legge & Foley, 1980). Both types of inhibition would decrease effective signal amplitude. Additive noise that occurs after the inhibitory stage (as in Bowne, 1990), would then decrease the effective signal-to-noise ratio. If the pooling stage follows the inhibitory and late additive noise stages, then the lower signal-to-noise ratio could offset the benefits of pooling and result in thresholds that do not change much with stimulus extent.

The mechanisms that might underlie our results, correlation and inhibition, could be due to feedforward connections from earlier stages relayed through lateral interactions, or be due to feedback from areas further along in the stream of visual processing. It is an open question as to which of these mechanisms or combinations thereof affect the early parsing of the visual image.

REFERENCES

Albright, T. D. (1984). Cortical processing of visual motion. In Miles, F. A. & Wallman, J. (Eds), *Visual motion and its role in the stabilization of gaze* (pp. 177–201). Amsterdam: Elsevier.

Albright, T. D. & Stoner, G. R. (1995). Visual motion perception. *Proceedings of the National Academy of Sciences USA, 92*, 2433–2440.

Allman, J., Miezin, F. & McGuinness, E. (1985). Direction- and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). *Perception, 14*, 105–126.

Bowne, S. F. (1990). Contrast discrimination cannot explain spatial frequency, orientation or temporal frequency discrimination. *Vision Research, 30*, 449–461.

Bravo, M. J. & Watamaniuk, S. N. J. (1995). Evidence for two speed signals: a coarse local signal for segregation and precise global signal for discrimination. *Vision Research, 35*, 1691–1697.

Britten, K. H., Shadlen, M. N., Newsome, W. T. & Movshon, J. A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. *Journal of Neuroscience, 12*, 4745–4765.

De Valois, R. L. & De Valois, K. K. (1990). *Spatial vision*. New York: Oxford University Press.

Downing, C. J. & Movshon, J. A. (1989). Spatial and temporal sensitivity in the detection of motion in stochastic random dot displays. *Investigative Ophthalmology and Visual Science, 29*, 72.

Engel, A. K., König, P. & Singer, W. (1991). Direct physiological evidence for scene segmentation by temporal coding. *Proceedings of the National Academy of Sciences USA, 88*, 9136–9140.

Graham, N., Kramer, P. & Yager, D. (1987). Signal-detection models for multidimensional stimuli: probability distributions and combination rules. *Journal of Mathematical Psychology, 31*, 366–409.

Gray, C. M., König, P., Engel, A. K. & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global properties. *Nature, 338*, 334–337.

He, Z. J. & Nakayama, K. (1992). Surfaces versus features in visual search. *Nature, 359*, 231–233.

He, Z. J. & Nakayama, K. (1994). Apparent motion determined by surface layout not by disparity or three-dimensional distance. *Nature, 367*, 173–175.

Hubel, D. H. & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology, 195*, 215–243.

Johnston, A. & Wright, M. J. (1986). Matching velocity in central and peripheral vision. *Vision Research, 26*, 1099–1109.

Legge, G. E. & Foley, J. M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America A, 70*, 1459–1470.

Maunsell, J. H. R. & Van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *Journal of Neurophysiology, 49*, 1127–1147.

McKee, S. P., Silverman, G. H. & Nakayama, K. (1986). Precise velocity discrimination despite random variations in temporal frequency and contrast. *Vision Research, 26*, 609–619.

Movshon, J. A., Adelson, E. H., Gizzi, M. S. & Newsome, W. T. (1985). The analysis of moving visual patterns. In Chagas, C., Gattass, R. & Gross, C. (Eds), *Pattern recognition mechanisms: Pontificiae Academiae Scientiarum Scripta Varia* (pp. 117–151). Rome: Vatican Press.

Movshon, J. A., Newsome, W. T., Gizzi, M. S. & Levitt, J. B. (1988). Spatio-temporal tuning and speed sensitivity in macaque visual cortical neurons. *Investigative Ophthalmology and Visual Science (Suppl.), 29*, 327.

Nakayama, K., He, Z. J. & Shimojo, S. (1996). Visual surface representation: a critical link between lower-level and higher-level vision. In Kosslyn, S. M. (Ed.), *Vision: Series invitation to cognitive science*. Cambridge, MA: MIT Press.

Newsome, W. T., Britten, K. H., Movshon, J. A. & Shadlen, M. (1989). Single neurons and the perception of visual motion. In Lam, D. M.-K. & Gilbert, C. (Eds), *Neural mechanisms of visual perception* (pp. 171–198). The Woodlands, TX: Portfolio.

Newsome, W. T. & Pare, E. B. (1988). Selective impairment of motion perception following ibotenic acid lesions of the middle temporal area of the macaque monkey. *Journal of Neuroscience, 8*, 2201–2211.

Pasternak, T. & Merigan, W. H. (1994). Motion perception following lesions of the superior temporal sulcus in the monkey. *Cerebral Cortex, 4*, 247–259.

Robson, J. G. & Graham, N. (1981). Probability summation and regional variation in contrast sensitivity across the visual field. *Vision Research, 21*, 409–418.

Rodman, H. R. & Albright, T. D. (1989). Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). *Experimental Brain Research, 75*, 53–64.

Salzman, C. D., Britten, K. H. & Newsome, W. T. (1990). Cortical stimulation influences perceptual judgements of motion direction. *Nature, 346*, 174–177.

Stoner, G. R. & Albright, T. D. (1992). Neural correlates of perceptual motion coherence. *Nature, 358*, 412–414.

Verghese, P. & Stone, L. S. (1995). Combining speed information across space. *Vision Research, 35*, 2811–2823.

Verghese, P. & Stone L. S. (1996). Perceived visual speed constrained by image segmentation. *Nature, 381*, 161–163.

Watson, A. B., Thompson, P., Murphy, B. J. & Nachmias, J. (1980). Summation and discrimination of gratings moving in opposite directions. *Vision Research, 20*, 341–347.

Welch, L. (1989). The perception of moving plaids reveals two motion-processing stages. *Nature, 337*, 734–736.

Welch, L. & Bowne, S. F. (1990). Coherence determines speed discrimination. *Perception, 19*, 425–435.

Zeki, S. M. (1974). Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *Journal of Physiology, 236*, 549–573.

Zipser, K. (1994). Sensitivity of V1 neurons to stereo-disparity information outside the receptive field. *Investigative Ophthalmology and Visual Science, 35*, 3171.

Acknowledgements—We thank Brent Beutter for helpful comments on the manuscript. This work was supported by NASA RTOP 199-16-12-37 to Lee Stone and a National Research Council Postdoctoral Associateship to Preeti Verghese.